

IN THE CLAIMS

1. - 34. (Cancelled)

35. (Currently Amended) An automated method for allocating resources among a plurality of resource-using computational entities in a data processing system, where the resources and the plurality of resource-using computational entities are hosted by a data center that provides computing services to a plurality of customers, the method comprising:

establishing a service-level utility for each of said plurality of resource-using computational entities, wherein the service-level utility is representative of an amount of business value obtained by each of said plurality of resource-using computational entities for one or more levels of performance and demand associated with each resource-using computational entity;

transforming said service-level utility into a resource-level utility for each of said plurality of resource-using computational entities, wherein the resource-level utility is representative of an amount of business value obtained by each of said plurality of resource-using computational entities when a quantity of said resources is allocated to the resource-using computational entity, wherein the resource-level utility indicates, for at least one of said plurality of resource-using computational entities, an estimated cumulative discounted or undiscounted future utility starting from current state descriptions of said at least one resource-using computational entity, wherein said estimated cumulative discounted or undiscounted future utility is trained on a temporal sequence of observed data using an adaptive machine learning procedure;

aggregating resource-level utilities of all of said plurality of resource-using computational entities;

computing a resource allocation from said resource-level utilities, as aggregated, by executing an optimization method to maximize a total utility of said data processing system, wherein said resource allocation involves re-allocating at least one of said resources from one of said plurality of resource-using computational entities to another

of said resource-using computational entities, wherein said optimization method comprises a standard linear or nonlinear algorithm, and wherein said computing further comprises computing a cost that is expected to be incurred as a result of said re-allocating; and

executing and conveying to the plurality of resource-using computational entities said resource allocation, where each of the plurality of resource-using computational entities uses said resource allocation to handle demands from the plurality of customers.